CourseProject2

Bamini Balaji

January 22, 2017

Part 2: Toothgrowth Data Analysis

Let us take a look at the toothgrowth data set:

```
# These are the first 5 lines of the dataset
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
# The dimensions of the data set:
dim(ToothGrowth)
```

```
## [1] 60  3
```

This data shows length of tooth under different doses and different types of supplements for 60 specimen.

```
summary(ToothGrowth)
```

```
##       len        supp         dose      
##  Min.   : 4.20   OJ:30   Min.   :0.500  
##  1st Qu.:13.07   VC:30   1st Qu.:0.500  
##  Median :19.25           Median :1.000  
##  Mean   :18.81           Mean   :1.167  
##  3rd Qu.:25.27           3rd Qu.:2.000  
##  Max.   :33.90           Max.   :2.000  
```

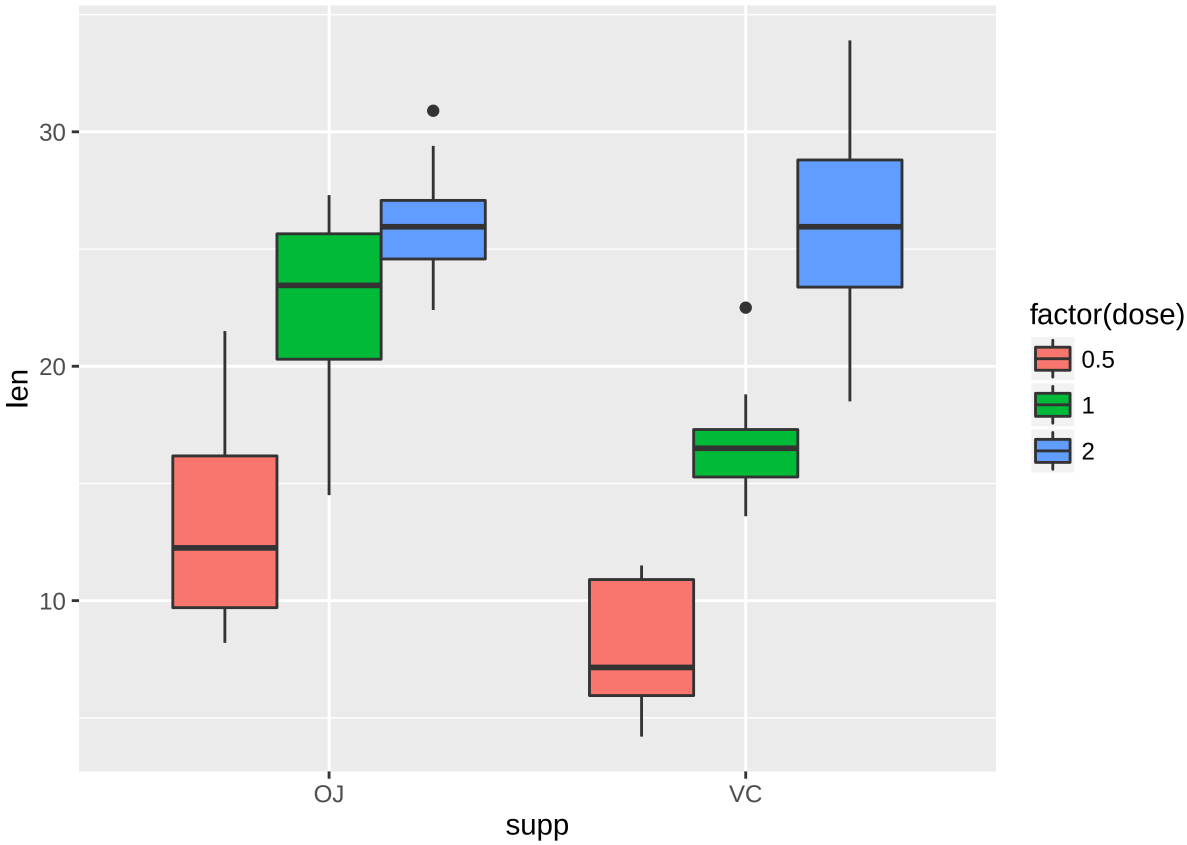
```
table(ToothGrowth$supp, ToothGrowth$dose)
```

```
##     
##      0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

The tooth length ranges from 4.20 to 33.90.

Supplements are either orange juice (OJ) or vitamin C (VC). Dose values are 0.5, 1, 2. For each combination of supplement and dose, 10 subjects are tested.

This figure shows the tooth length (y-axis) versus supplement OJ or VC for the various dose levels.



OJ versus VC

Does orange juice (OJ) stimulate tooth growth over vitamic C (VC)?

```
with(ToothGrowth, t.test(len[supp == "VC"], len[supp == "OJ"]))
```

```
##
##  Welch Two Sample t-test
##
## data:  len[supp == "VC"] and len[supp == "OJ"]
## t = -1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.5710156  0.1710156
## sample estimates:
## mean of x mean of y
##  16.96333  20.66333
```

```
pval <- with(ToothGrowth, t.test(len[supp == "VC"], len[supp == "OJ"]))$p.value
```

p-value > alpha. Hence null hypothesis that VC and OJ are equivalent is accepted

Let's look at the data for a specific dose level of 0.5

```r
with(ToothGrowth, t.test(len[supp == "VC" & dose == 0.5], len[supp == "OJ" & dose == 0.5]))
```

```
##
##  Welch Two Sample t-test
##
## data:  len[supp == "VC" & dose == 0.5] and len[supp == "OJ" & dose == 0.5]
## t = -3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.780943 -1.719057
## sample estimates:
## mean of x mean of y
##      7.98     13.23
```

```r
pval <- c(pval, with(ToothGrowth, t.test(len[supp == "VC" & dose == 0.5], len[supp == "OJ" & dose == 0.
```

Here p-value < alpha. The confidence interval does not encompass 0. Hence alternative hypothesis is accepted where the OJ and VC have different impact at 0.5 dose. In fact, since the values are negative in confidence interval range, it suggests that OJ > VC for 0.5 dose.

Now, let's look at dose of 2.

```r
with(ToothGrowth, t.test(len[supp == "VC" & dose == 2], len[supp == "OJ" & dose == 2]))
```

```
##
##  Welch Two Sample t-test
##
## data:  len[supp == "VC" & dose == 2] and len[supp == "OJ" & dose == 2]
## t = 0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.63807  3.79807
## sample estimates:
## mean of x mean of y
##     26.14     26.06
```

```r
pval <- c(pval, with(ToothGrowth, t.test(len[supp == "VC" & dose == 2], len[supp == "OJ" & dose == 2]))$
```

In this case the null hypothesis is accepted. This suggests that at high doses, OJ and VC are equivalent.

Effect of dose level

Let's compare a dose of 2 and a dose of 1 collectively regardless of supplement type.

```r
with(ToothGrowth, t.test(len[dose == 2], len[dose == 1]))
```

```
##
##  Welch Two Sample t-test
##
## data:  len[dose == 2] and len[dose == 1]
```

```
## t = 4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.733519 8.996481
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```

```r
pval <- c(pval, with(ToothGrowth, t.test(len[dose == 2], len[dose == 1]))$p.value)
```

The data confirms that our null hypothesis is rejected. Dose level of 2 causes greater tooth length than dose level 1.

What about in the case of OJ at dose levels 2 and 1. These two data appear close.

```r
with(ToothGrowth, t.test(len[supp == "OJ" & dose == 2], len[supp == "OJ" & dose == 1]))
```

```
##
##  Welch Two Sample t-test
##
## data:  len[supp == "OJ" & dose == 2] and len[supp == "OJ" & dose == 1]
## t = 2.2478, df = 15.842, p-value = 0.0392
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1885575 6.5314425
## sample estimates:
## mean of x mean of y
##     26.06     22.70
```

```r
pval <- c(pval, with(ToothGrowth, t.test(len[supp == "OJ" & dose == 2], len[supp == "OJ" & dose == 1]))$
```

In this instance as well, the data shows p-value > alpha. Hence null hypothesis is rejected suggesting these two groups are not equivalent.

Overall these tests suggest that dose level has a large impact. At low doses, OJ is more effective than VC. But at high doses this effect does not matter.

Multiple Testing

Since we have used this data for multiple testing, it would be advisable to avoid false discoveries. This can be done my using the BH correction.

```r
pval
```

```
## [1] 0.0606345079 0.0063586068 0.9638515887 0.0000190643 0.0391951420
```

```r
p.adjust(pval, method = "BH")
```

```
## [1] 7.579313e-02 1.589652e-02 9.638516e-01 9.532148e-05 6.532524e-02
```

The last test comparing tooth length with OJ at dose level 2 versus 1 is no longer showing a significant finding since p-value > alpha. Hence these two conditions would be considered equivalent.